Asymptotic Analysis

Asymptotic Analysis Definitions

Let f, g be functions from the positive integers to the non-negative reals.

Definition 1: (Big-Oh notation)

$f = O(g)$ if there exist constants $c > 0$ and n_0 such that for all $n > n_0$,

$$f(n) \leq c \cdot g(n).$$

Definition 2: (Big-Omega notation)

$f = \Omega(g)$ if there exist constants $c > 0$ and n_0 such that for all $n > n_0$,

$$f(n) \geq c \cdot g(n).$$

Definition 3: (Big-Theta notation)

$f = \Theta(g)$ if $f = O(g)$ and $f = \Omega(g)$.

Note: You will use "Big-Oh notation", "Big-Omega notation", and "Big-Theta notation" A LOT in class. Additionally, you may occasionally run into "little-oh notation" and "little-omega notation". You are not responsible for knowing the following definitions in this class:

Definition 4:(Little-oh notation)

$f = o(g)$ if **for every constant** $c > 0$ there exist a constant n_0 such that for all $n > n_0$,

$$f(n) < c \cdot g(n).$$

Definition 5:(Little-omega notation)

$f = \omega(g)$ if **for every constant** $c > 0$ there exist a constant n_0 such that for all $n > n_0$,

$$f(n) > c \cdot g(n).$$

Asymptotic Analysis Problems

1. Prove that if $f = \Omega(g)$ then f is not in $o(g)$.

2. For each of the following functions, prove whether $f = O(g)$, $f = \Omega(g)$, or $f = \Theta(g)$. For example, by specifying some explicit constants n_0 and $c > 0$ such that the definition of Big-Oh, Big-Omega, or Big-Theta is satisfied.

(a)	$f(n) = n\log(n^3)$	$g(n) = n\log n$
(b)	$f(n) = 2^{2n}$	$g(n) = 3^n$
(c)	$f(n) = \displaystyle\sum_{i=1}^{n} \log i$	$g(n) = n\log n$

3. Give an example of f, g such that f is not $O(g)$ and g is not $O(f)$.

Divide & Conquer

4. In lecture, you have seen how digit multiplication can be improved upon with divide and conquer. Let us see a more generalized example of Matrix multiplication. Assume that we have matrices A and B and we'd like to multiply them. Both matrices have n rows and n columns.

   For this question, you can make the simplifying assumption that the product of any two entries from A and B can be calculated in O(1) time.

   (a) What is the naive solution and what is its runtime? Think about how you multiply matrices.

   (b) Now if we divide up the problem like this:

   $$XY = \begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} E & F \\ G & H \end{bmatrix} = \begin{bmatrix} AE + BG & AF + BH \\ CE + DG & CF + DH \end{bmatrix}$$

   We now have a divide and conquer strategy! Find the recurrence relation of this strategy and the runtime of this algorithm.

   (c) Can we do better? It turns out we can by calculating only 7 of the sub problems:

$P_1 = A(F - H)$	$P_5 = (A + D)(E + H)$
$P_2 = (A + B)H$	$P_6 = (B - D)(G + H)$
$P_3 = (C + D)E$	$P_7 = (A - C)(E + F)$
$P_4 = D(G - E)$	

   And we can solve XY by

   $$XY = \begin{bmatrix} P_5 + P_4 - P_2 + P_6 & P_1 + P_2 \\ P_3 + P_4 & P_1 + P_5 - P_3 - P_7 \end{bmatrix}$$

   We now have a more efficient divide and conquer strategy! What is the recurrence relation of this strategy and what is the runtime of this algorithm?

How NOT to prove claims by induction

5. In this class, you will prove a lot of claims, many of them by induction. You might also prove some wrong claims, and catching those mistakes will be an important skill!

   The following is an example of a false proof where an obviously untrue claim has been 'proven' using induction (with some errors or missing details, of course). Your task is to investigate the 'proof' and identify the mistake made.

   Fake Claim 2:

   $$\underbrace{\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \cdots}_{n \text{ terms}} = \frac{3}{2} - \frac{1}{n}. \tag{1}$$

   Inductive Hypothesis: (1) holds for $n = k$

   Base Case: For $n = 1$,

   $$\frac{1}{1 \cdot 2} = 1/2 = \frac{3}{2} - \frac{1}{1}.$$

   Inductive Step: Suppose the inductive hypothesis holds for $n = k$; we will show that it is also true $n = k + 1$. We have

   $$\left(\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \cdots \frac{1}{(k-1) \cdot k} \right) + \frac{1}{k \cdot (k+1)} = \frac{3}{2} - \frac{1}{k} + \frac{1}{k \cdot (k+1)} \quad \text{(by weak induction hypothesis)}$$
   $$= \frac{3}{2} - \frac{1}{k} + \frac{1}{k} - \frac{1}{k+1}$$
   $$= \frac{3}{2} - \frac{1}{k+1}.$$

   Conclusion: By weak induction, the claim follows.

Weak vs. Strong Induction

The difference between these two types of inductions appears in the inductive hypothesis. In weak induction, we only assume that our claim holds at the k-th step, whereas in strong induction we assume that it holds at all steps from the base case to the k-th step. In this section, let's examine how the two strategies compare.

6. Consider the following proof by weak induction.

   Claim: For any positive integer n, $6^m - 1$ is divisible by 5.

   Inductive Hypothesis: The claim holds for $n = k$. I.e., $6^k - 1$ is divisible by 5,

   Base Case: For $n = 1$,

   $$6^n - 1 = 5 = 5(1)$$

Inductive Step: Suppose the inductive hypothesis holds for $n = k$; we will show that it is also true $n = k + 1$. We have

$$6^{k+1} - 1 = 6(6^k) - 1$$
$$= 6(6^k - 1) - 1 + 6$$
$$= 6(6^k - 1) + 5$$

By the weak inductive hypothesis, $6(6^k - 1)$ is divisible by 5, and the second term is also clearly divisible by 5. Therefore, $6^{k+1} - 1$ is divisible by 5.

Conclusion: By weak induction, the claim follows.

Would it be sufficient to use strong induction instead of weak induction for this proof?

7. Now consider the following proof by strong induction.

First, we introduce a game called Nim, in which there are two players and two piles of matches. At each turn, a player removes some (non-zero) number of matches from one of the piles. The player who removes the last match wins.

Consider the winning strategy for the second player: Suppose your opponent removes m matches from one pile. You remove m matches from the other pile. We prove by induction that this strategy is in fact a winning strategy by inducting on the number of matches in each pile.

Claim: If the two piles contain the same number of matches n at the start of the game, then by using the above strategy, the second player will always win.

Inductive Hypothesis: The claim holds for all values n where $1 \leq n \leq k$. I.e., if both piles have n matches at the beginning of the game, then the second player can always win.

Base Case: When there are exactly 1 matches in both piles, then the first player is forced to pick 1 match from one of the piles. Player 2 can win by picking 1 match from the other pile.

Inductive Step: Suppose the inductive hypothesis holds for $1 \leq n \leq k$, we will show that it also holds for $n = k + 1$.

If both piles contain $k + 1$ matches at the beginning of the game, any legal move by the first player involves removing j matches from one pile, where $0 \leq j \leq k+1$. The piles then contain $k + 1$ matches and $k + 1 - j$ matches.

The second player can now remove j matches from the other pile. This leaves us with two piles of $k + 1 - j$ matches. If $j = k + 1$, then the second player wins. If $j < k + 1$, we can use the inductive hypothesis to conclude that the second player can win from this new starting point in the game.

Conclusion: By strong induction, the claim follows.

Would it be sufficient to use weak induction instead of strong induction here?

More Induction

8. On a flat ice sheet, an *odd* number of penguins are standing such that their pairwise distances to each other are all different. At the strike of dawn, each penguin throws a snowball at another penguin that closest to them. Show that there is always some penguin that doesn't get hit by a snowball.